1a Ethical sourcing of data. Data that has been procured must have a valid indication of consent and traceable source for audit. Exceptions for public use maybe be found in copyright law. It is notable that use of any data is permanent i.e. it is impossible for "unlearn" what it has learned from using any data. Therefore if copyrighted images, for example, are used without permission, the AI 'Maker' must reset the whole AI and restart training from square 1 to be confident in having removed the data.

1d AI audits should be folded into other accountability mechanisms. For example, Use of copyright images or Use of likeness can fall under Intellectual Property protections. Other regulations like regulations for diversity should only apply where if the task were performed traditionally, relevant concerns apply. Hiring is one example that comes to mind.

1e There is dangerous amount of potential for impacts of legal standards. For things like unauthorized use of copyright images in training AI, A developer must delete the whole model and start anew or load a backup save from before the copyrighted image was used to be sure to have removed the copyright image from the AI. There is no such thing as "unlearning" data. It is like accidentally having taught and trained a citizen the manufacture of illicit drugs and arms and then telling them to forget that they have even learned it. Even if they tried, an inkling of those skills will remain. Courts must be sure and consistent in their interpretation of existing laws that manage information. Intellectual Property, Pornography Regulation, Medical Information Protection, Consent and many more topics are relevant and must be used to scrutinize sources of data. Legislatures must then make clear what the standards are in regards to data used in AI.

2 Certification and thorough audits will build trust but trust is earned by actually conducting business ethically. It is pointless to play charades to a knowledgeable skeptic. Transparency and publicly reviewable provenance to data used will incredibly increase trust from a lay-consumer. Stakeholders may feel regulations are useless and will incur unnecessary cost. Policy in this category do not specifically protect the consumer nor the stakeholder, rather they protect the safety of any prospective data source. That can be anyone who generates information online or offline. Pictures, written letters, digital art, tweets are all examples of individual generated information.

3d like with the tiktok hearing, without competent and qualified legislators, i.e. those who have received a higher education degree in the field, No amount of legislation will be valid or competent.

3g With the current lack of transparency, consultation or redress is impossible. It seems to not have been considered in the first place.

4 No. AIs depend on historic data. Historic data contain issues that current non AI systems generate.

5 Considering current generative models have started without regulation, it is near impossible to retrofit current models to be compliant. There is only extensive training that will only ensure a high probability for, but not certain compliance.

7 Like the gun debate, AI is a tool just as the gun is a tool; it is how it is used that is the issue. Regulating hardware or training rate will only drive black market economies to work around regulation. To regulate AI, regulating data sourcing and credentialing data sourcing companies or operations is the best way to limit harm done by AI for data dictates the form of the model. Form dictates function.

9 Currently, It depends on the moral compasses of the host company and existing laws that may drive company policy. There are no external sectors, industries, or specialist companies that stand above the rest, although larger companies do tend to run up more ethics scandals due to an overwhelming need to pursue profit as fast as possible.

10 Fair, safe, effective, and trustworthy are all subjective terms. They vary from individual to individual. When conversing officially with another entity that uses such terms it is imperative to demand their definitions of such words written with an mutually agreed definition. For example a lot of nonconsensual pornographic deepfakes are considered safe by their hosts and creators but dangerous by activists and some subjects of said deepfakes.

11 All areas have their shortfalls. Borrowing from the finance sector and activist investing, however, the profits aren't just the only concern, the provenance of those profits, i.e. how they were generated matters too. Same goes with sourced data. It is not just that one has data, one must log where it came from and if the data is authorized to be used in an AI.

13 a right to privacy is imperative. Any data collection should be an opt-in

mechanism. Those in government must be accredited in the field to be qualified to preside over conflicts or crimes regarding these issues. A lawyer by trade would be hard pressed to understand how a rocket flies on a fundamental level. Therefore a government official that wishes to legislate, judge, or enforce on any issue must have been trained to understand and apply relevant concepts in accordance to policy.

15a Data sourcing. Sourced data must have consent to be used in training of models.

15b If data generated downstream of the model runs afoul of regulations, then it is imperative to comb through the source of that data. A baby is hard pressed to identify a word or sentence if it has not been taught any language.

15c Credible Audits should always be conducted by an unassociated third party to prevent conflict of interest, just as officers, doctors, judges and the like are required to recuse themselves on the case if they are likely to be personally invested in a case or issue at hand.

20 As said before, Sources and Consent to use specific pieces of data should be kept. Logs of potential compliance issues and actions taken to remedy that as well as the remedy result should be logged. AI is like a child. If you swear in front of it, it will swear right back at you. Therefore a developer MUST strictly control inflow of information and monitor outflow for signs of non compliance. There is no accountability by design except for extensive logging of any change be it actions within the model, actions outside the model as well as the inflow and outflow of data.

21 Data sources MUST be traceable. To use data without knowing the source is like driving a car because "I found the keys still in the car and took the car." Regardless, logs like debugging or general operational logs should be kept for forensic review.

22 Data curation is a paid service. To force equal access to already curated data is unfair to data curators. However, Data curation should be licensed like a medical license. If any person passes a regulated bar and renews their license accordingly then it is reasonable that any person who wishes to have access to data can learn how to curate data and be licensed to to it for themselves or as a paid service. A more apt analogy should be the driver's license but that IS putting too much faith in the general population and corporations alike. For an audited company to gain data, they must employ properly licensed data curators to filter through and curate data that is ethically sourced and on the subject matter. Professionals that procure data from dubious sources should be in violation of their license i.e. committed malpractice.

24 Independent auditing companies should be formed and regularly inspected to be unassociated with and if associated, forcefully disassociated or disbanded. if unions form, they must also be separated along the lines of auditors and the audited.

25 YES. DATA PROTECTION, DATA PRIVACY, AND INTELLECTUAL PROPERTY PROTECTION LAWS ARE CURRENTLY SEVERELY LACKING AND HAVE CAUSED AND WILL CAUSE OPPORTUNITY FOR UNAUTHORIZED USE OF DATA BY POLITICALLY STRONG ORGANIZATIONS. THERE IS RAMPANT DATA COLLECTION AND LITTLE TRANSPARENCY. If, for example, one requests google to delete their activity logs from their servers, Google may say it is deleted but there is no way to confirm that status. I.e. they have no way of knowing that the deletion mechanism is functioning as intended, whether or not there are offline copies of their data that is impervious to automatic deletion, etc. Further if that data is already used in training of any AI model, that data is permanently stored implicitly by the arrangement and weights of neurons within a neural network (a.k.a. an AI model).

26 There is a lack of federal law that dictates who in the legislative system is qualified to legislate on the matter. The recent tiktok hearing has drawn international ridicule over the senate's technological competency. The most common joke is that a member of congress's median technological skill is determined by the age of their children; Most have to ask their adolescent children to gain knowledge of even fundamental items like the smartphone much less cryptographic techniques or software surveillance. Other issues like corruption has been highlighted by Clarence Thomas's scandal. For an official to be in such a historically sacred nexus of political power and morality and had been actively receiving perceived benefits from a political lobbyist can and should call into question the corruption levels and ethics of all lower offices. In my personal opinion, politicians should never have been able to receive political donations in the first place. Policy is not and should never be a business.

27 Nondisclosure and trade secrets on the model's inner workings is alright, in my personal opinion. However, it should be readily available to any layman

any aspect of the input data such as source and consent for use. If public access to the raw data constitutes an economic loss, then a third party licensed "describer" or generally an auditor may create a text label to paint in broad but accurate strokes the general nature of the data. Source and consent must be fully visible however.

30a A keystone regulation should be intensive and robust licensing of data curators including testing for a license. This should be implemented like a medical license, food handler's card, driver's license, or any similar licensing.

30b Input data including provenance and expressed consent for use for all used data should be readily provided. Logs of noncompliance, remedial actions for said noncompliance, and results of said remedial actions should be readily available. When noncompliance happens, Two copies of noncompliance logs should be generated, one copy is to be stored securely by the company or entity in question and the other should be stored by a specialized government regulatory body regarding AI.

30c Any person collecting data for commercial or even public use must be first be a licensed data curator. Licensing should also be handled at a national level due to the borderless nature of the internet. Currently existing agencies aren't equipped to test citizens and issue licenses but investigative bodies like the FBI seem more than well equipped to enforce regulations.

30d There should be independent watchdog organizations consisting of officers that are accredited professionally in the field.

31 What specific activities should government fund to advance a strong AI accountability ecosystem? Any activity in particular. The government must develop and inform itself on dangers and effective remedies to AI issues like how governments already keep level 4 biohazard labs that study possible super-contagions.